

Lesley E. Weaver (SBN 191305)
Anne K. Davis (SBN 267909)
Joshua D. Samra (SBN 313050)
BLEICHMAR FONTI & AULD LLP
1330 Broadway, Suite 630
Oakland, CA 94612
Tel. (415) 445-4003
lweaver@bfalaw.com
adavis@bfalaw.com
jsamra@bfalaw.com

Gregory S. Mullens (admitted *pro hac vice*)
BLEICHMAR FONTI & AULD LLP
75 Virginia Road, 2nd Floor
White Plains, NY 10603
Tel. (415) 445-4006
gmullens@bfalaw.com

Plaintiffs' Interim Co-Lead Counsel

**UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA
SAN JOSE DIVISION**

In re Google Generative AI Copyright Litigation

Joseph R. Saveri (SBN 130064)
Cadio Zirpoli (SBN 179108)
Christopher K.L. Young (SBN 318371)
Elissa A. Buchanan (SBN 249996)
Evan A. Creutz (SBN 349728)
Aaron Cera (SBN 351163)
JOSEPH SAVERI LAW FIRM, LLP
601 California Street, Suite 1505
San Francisco, CA 94108
Telephone: (415) 500-6800
Facsimile: (415) 395-9940
Email: jsaveri@saverilawfirm.com
czirpoli@saverilawfirm.com
cyoung@saverilawfirm.com
eabuchanan@saverilawfirm.com
ecreutz@saverilawfirm.com
acera@saverilawfirm.com

Master File Case No. 5:23-cv-03440-EKL-SVK
Consolidated Case No. 5:24-cv-02531-EKL-SVK

**DECLARATION OF STEPHEN J. TETI IN
SUPPORT OF PLAINTIFFS' MOTION TO
AMEND CASE SCHEDULE**

UNSEALED DOCUMENT (ECF NO. 224-3)

1 I, Stephen J. Teti, hereby declare as follows:

2 1. I am admitted *pro hac vice* before this Court and a Partner with Lockridge Grindal Nauen
3 PLLP, Counsel for Plaintiffs in this matter. I have personal knowledge of the facts stated herein and, if
4 called upon to do so, could and would testify competently thereto. I submit this Declaration in support of
5 Plaintiffs' Motion under Fed. R. Civ. P. 16, Civil Local Rule 16-10, and this Court's inherent authority, to
6 Amend the Case Schedule. I have personal knowledge of the facts stated herein and, if called upon to do
7 so, could and would testify competently thereto.

8 2. Plaintiffs did not gain access to Google's training data during the week of June 23, 2025,
9 although the Court had ordered that. ECF No. 155. Google failed to timely provision user credentials or
10 provide adequate instructions to set up the Google-provided Chromebooks required to access the remote
11 environment. Plaintiffs and their experts did not gain access to the training data environment until July 3,
12 2025.

13 3. Google did not make its technical personnel available for conferral with Plaintiffs' experts
14 until July 29, 2025.

15 4. Throughout July, Plaintiffs conferred with Google because Plaintiffs' experts could not
16 analyze the data Google produced in its environment without enabling certain analytical tools.

17 5. On a July 29, 2025 call with Google's engineers—the first time Plaintiffs' experts were
18 allowed to confer with Google's technical support, and only after a court order—Plaintiffs learned for the
19 first time from Google's technical engineers that Google did not use Google's BigQuery database tables in
20 their chosen environment. Google's engineers advised that Plaintiffs instead must access training data via
21 TensorFlow, and that Google would need to convert relevant tables to TFRecord format. MK Analytics
22 was not provided with access to TensorFlow until August 9. *See* Declaration of Meredith "Merry"
23 McCarron ("McCarron Decl."), dated September 12, 2025, attached hereto as **Exhibit 1**.

24 6. On August 13, 2025, Plaintiffs' experts, MK Analytics, notified me that 31 datasets that
25 Google had converted to TFRecord were either corrupted or inaccessible. I notified Google that same day
26 of the problem. Google responded on August 15, 2025, acknowledged that there were "hundreds of
27 thousands" of file locations, and sought "clarification." After Plaintiffs responded to the best of their ability
28

1 on August 18, on August 21, 2025, Google provided a substantive response to my email of August 13,
2 2025.

3 7. On Saturday, August 30, 2025, Google revealed to Plaintiffs for the first time that critical
4 metadata, including URL information from which the data was crawled, was not provided in the training
5 datasets converted to TFRecord. I responded to Google's email the next day, August 31, 2025, requesting
6 additional information.

7 8. On Tuesday, September 2, 2025, I received an email from Google's counsel, Ms. Knoll,
8 informing us that the metadata error affected a total of 49 datasets (or "files" as referred to by Google),
9 including the 31 datasets referenced above. *See* McCarron Decl. ¶ 21. Ms. Knoll indicated that Google
10 expected to complete conversions by the next day.

11 9. During a meet and confer on September 3, 2025, Google's counsel disclosed for the first
12 time that Google had discovered errors in its conversion of the training data during the week of August 19,
13 2025, but had not informed Plaintiffs of these issues until August 30, 2025.

14 10. During the same September 3, 2025 conference, Google's counsel admitted that Google had
15 unilaterally added to and modified some of the training data in the review environment without informing
16 Plaintiffs first, while Plaintiffs' experts were actively conducting confidential analysis.

17 11. During the same September 3, 2025 conference, Google also disclosed for the first time that
18 additional metadata fields such as DocIDs may also have been altered or omitted as a result of the
19 conversion. When Plaintiffs' counsel asked Google's counsel during the September 3, 2025 conference
20 about what other metadata beyond URLs and DocIDs may have been affected and for which models,
21 Google was unable to answer.

22 12. With no knowledge that the datasets were incomplete, Plaintiffs' experts had been working
23 in the Google environment, attempting to understand recurring issues.

24 13. Recently, when Plaintiffs asked Google to restore one of our expert's access to the training
25 data environment, Google responded that "[r]esets to the system can only be performed during work hours,
26 and for security reasons require two Google personnel to be available."

27 14. On September 9, 2025, Google provided to Plaintiffs two "takeout" files
28 (gemini_v1_books_parsed_concat.pkl and lamda_books_parsed_concat.pkl) representing portions of the

1 training data identified by Plaintiffs' experts for analysis outside of the training data review environment.
2 Plaintiffs identified these files for export from the training data review environment on August 28, 2025,
3 triggering a five-business day deadline for their delivery pursuant to an agreed-upon provision of the
4 training data protocol and the Parties' prior course of conduct for Plaintiffs' requested takeouts.

5 15. That same day, Google informed Plaintiffs that it would not produce a third takeout file
6 (lamda_c4_docjoins_parsed_concat.pkl), which Plaintiffs had identified to Google on August 29, 2025,
7 because Google considered the size of the takeout file to be too large, at 600 GBs.

8 16. I understand that the training data environment contains petabytes of data, such that a 600
9 GB takeout represents a tiny fraction of the data as a whole, and consists of a subset of data relevant to
10 Plaintiffs' class certification motion. *See* McCarron Decl. ¶ 14.

11 17. The protocol governing access to Google's training data was submitted to the Court on June
12 20, 2024. ECF No. 159-1. Plaintiffs have been operating pursuant to its terms since that date. No express
13 size limitations are reflected in the protocol governing training data, nor did Google raise any size
14 limitations as to the amount of takeout data Plaintiffs are entitled to request, until September 9, 2025.

15 18. Google withdrew its objections to Plaintiffs' takeout requests on September 11, 2025,
16 approximately one hour before a scheduled discovery hearing before Magistrate Judge van Keulen. Judge
17 van Keulen ordered that the time for Google to comply with Plaintiffs' takeout requests be narrowed from
18 five to three business days. *See* ECF No. 221. Google produced the previously-disputed takeout on
19 September 12, 2025.

20 19. On September 12, 2025, Google disclosed for the first time that there is a file-number limit
21 for takeouts. This has postponed production of another set of Plaintiffs' takeout requests.

22 20. Plaintiffs met and conferred with counsel for Google several times about the subject matter
23 of this Motion, including most recently on September 12, 2025. Google would not agree to more than a
24 three-week extension, which proposal would not also extend any deadlines following after the class
25 certification hearing.

26 21. Attached hereto as **Exhibit 2** is a true and correct copy of the Declaration of Ayyub Ibrahim,
27 dated September 12, 2025.
28

1 I declare under penalty of perjury under the laws of the United States of America that the foregoing
2 is true and correct. Executed this 12th day of September, 2025, at Mansfield, Massachusetts.

3 /s/ Stephen J. Teti
4 Stephen J. Teti
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

ATTESTATION PURSUANT TO CIVIL LOCAL RULE 5-1(i)(3)

I, Lesley E. Weaver, attest that concurrence in the filing of this document has been obtained from the other signatories. I declare under penalty of perjury that the foregoing is true and correct.

Executed this 12th day of September, 2025, at Oakland, California.

/s/ Lesley E. Weaver

Lesley E. Weaver